

Incorporating Asymmetric Distributional Information in Robust Value-at-Risk Optimization

Karthik Natarajan

Singapore-MIT Alliance and Department of Mathematics, National University of Singapore, Singapore 117543,
matkbn@nus.edu.sg

Dessislava Pachamanova

Mathematics and Sciences Division, Babson College, Babson Park, Massachusetts 02457,
dpachamanova@babson.edu

Melvyn Sim

Singapore-MIT Alliance and NUS Business School, National University of Singapore, Singapore 117592,
dscsimm@nus.edu.sg

Value-at-Risk (VaR) is one of the most widely accepted risk measures in the financial and insurance industries, yet efficient optimization of VaR remains a very difficult problem. We propose a computationally tractable approximation method for minimizing the VaR of a portfolio based on robust optimization techniques. The method results in the optimization of a modified VaR measure, Asymmetry-Robust VaR (ARVaR), that takes into consideration asymmetries in the distributions of returns and is coherent, which makes it desirable from a financial theory perspective. We show that ARVaR approximates the Conditional VaR of the portfolio as well. Numerical experiments with simulated and real market data indicate that the proposed approach results in lower realized portfolio VaR, better efficient frontier, and lower maximum realized portfolio loss than alternative approaches for quantile-based portfolio risk minimization.

Key words: value-at-risk; robust optimization; coherent risk measures

History: Accepted by Dorit Hochbaum, optimization and modeling; received August 15, 2005. This paper was with the authors 2 months for 3 revisions. Published online in *Articles in Advance* January 14, 2008.

1. Introduction

Markowitz (1952) made a substantial contribution to portfolio theory by casting the issue of “best” portfolio allocation as an optimization problem. He suggested finding an asset allocation that results in the minimum portfolio risk, as represented by the portfolio variance, for a given level of target portfolio expected return. Portfolio risk management theory has made significant progress since Markowitz’s seminal paper. It is now well known that while mean-variance optimization is appropriate for symmetrically distributed portfolio returns, it results in unsatisfactory asset allocations when returns are asymmetrically distributed, or when downside risk is more weighted than upside risk. Such considerations and the theory of stochastic dominance (Levy 1992) have spurred interest in asymmetric or quantile-based risk measures, e.g., semivariance, Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), and others (Jorion 2000, Dowd 1998, Konno and Yamazaki 1991, Cariño and Turner 1998, Rockafellar and Uryasev 2002). Among those, VaR remains the most widely accepted measure among practitioners. VaR estimates the maximum potential loss at a certain probability level, i.e., it provides information about the amount of losses that

will not be exceeded with certain probability. Mathematically, $(1 - \epsilon)$ -VaR is defined as the minimum level γ_0 such that the probability that the portfolio loss exceeds γ_0 is below ϵ , where $\epsilon \in (0, 1]$ is the probability level specified by the user (typically, $\epsilon = 1\%$ or 5%). Thus, the portfolio VaR optimization problem with probabilistic constraint can be formulated as

$$\begin{aligned} \min \quad & \gamma_0 \\ \text{s.t.} \quad & P(V_0 - \tilde{\mathbf{r}}'\mathbf{x} > \gamma_0) \leq \epsilon, \\ & \mathbf{x}'\mathbf{e} = 1, \end{aligned} \quad (1)$$

where V_0 is the value of the portfolio today, $\mathbf{x} \in \mathcal{R}^N$ is the vector of asset weights, and $\tilde{\mathbf{r}} \in \mathcal{R}^N$ is the vector of (uncertain) future portfolio asset returns, i.e., $\tilde{\mathbf{r}}'\mathbf{x}$ is the value of the portfolio at a predetermined date in the future. $V_0 - \tilde{\mathbf{r}}'\mathbf{x}$ is the change in portfolio value between today and this predetermined date. The optimization problem can be reformulated as

$$\begin{aligned} \min \quad & \gamma \\ \text{s.t.} \quad & P(\gamma + \tilde{\mathbf{r}}'\mathbf{x} \geq 0) \geq 1 - \epsilon, \\ & \mathbf{x}'\mathbf{e} = 1, \end{aligned} \quad (2)$$

where $\gamma = \gamma_0 - V_0$.

Despite VaR's popularity, there are some major problems with using it as the risk measure in portfolio optimization. First, VaR is not a coherent risk measure—it lacks subadditivity in the sense of Artzner et al. (1997). Thus, the VaR of a portfolio of two funds may end up greater than the sum of the VaRs of the individual funds. Second, even though the purpose of VaR is to reduce extreme losses, minimization of VaR can lead to an undesirable stretch of the tail of the distribution exceeding VaR (Rockafellar and Uryasev 2000). Third, VaR optimization is a stochastic programming problem of a special kind (namely, one with probability constraints), and as such, it is inherently difficult to solve. Indeed, while a substantial part of the VaR estimation literature deals with the problem of computing the VaR given a particular asset allocation, the literature on actual optimization of portfolio VaR is relatively scarce (see <http://www.gloriamundi.org> for links to a number of papers on methodologies for modeling VaR).

It is important to emphasize that unless the probability distributions of future returns are known exactly, the actual value of the portfolio VaR cannot be obtained. Thus, in practice all methods for optimization of VaR resort to approximations, either by applying optimization directly to a sample that approximates the actual portfolio loss distribution, or by specifying a particular shape for the return distributions, and computing important distributional parameters from data. In industry, as well as in a large part of the VaR literature, it is frequently assumed that returns are normal (Dowd 1998, Jorion 2000). Under this assumption, VaR is a scalar multiple of the standard deviation, and the portfolio optimization problem for VaR can be formulated similarly to the traditional Markowitz (1952) problem. Theoretically, this assumption is not unreasonable because the Central Limit Theorem implies that over the long horizon, stock returns should be approximately Gaussian as long as short-horizon returns are sufficiently independent (see, for example, Campbell et al. 1997). Empirically, however, there is evidence that both short- and long-horizon stock returns can be skewed and highly leptokurtic (Fama 1976, Duffee 2002). Furthermore, the returns of portfolios involving derivatives or credit risky assets can have extremely left-skewed distributions (Schönbucher 2000). Thus, the normality assumption may lead to gross underestimation of the actual portfolio VaR.

As an optimization problem with chance constraints, portfolio VaR management can be approached with tools from *robust optimization*, which has become a leading methodology for dealing with uncertainty in optimization problems in recent years. El Ghaoui et al. (2003) used robust optimization to propose a framework for optimization of worst-case VaR based

on information about first and second moments of the distribution of returns. We build on their idea by suggesting a framework for approximate VaR optimization based on more general asymmetric measures of variability for the distribution of returns, following results from Chen et al. (2007). Our contributions can be summarized as follows:

(a) We generalize the framework of El Ghaoui et al. (2003) by suggesting an approach to VaR optimization that reduces to their approach for particular values of the parameters, but is also capable of incorporating information about asymmetries in the distributions of uncertainties.

(b) We prove that the modified VaR we obtain (we name it “Asymmetry-Robust VaR,” or ARVaR) not only closely approximates the actual VaR, but is also a coherent risk measure in the sense of Artzner et al. (1997). Coherence has become an important consideration in risk management in recent years, as illustrated by the increase in popularity of the coherent quantile-based risk measure CVaR in the insurance and the financial industries (Rockafellar and Uryasev 2000, 2002).

(c) We note that ARVaR approximates CVaR. Direct optimization of CVaR is a convex optimization problem; however, it involves the evaluation of a multidimensional integral, and is computationally prohibitive in the sense that it is not necessarily polynomial-time solvable in the input size of the problem. *Sample* CVaR optimization is a linear programming problem, and as such is more attractive computationally; however, its performance can be very sensitive to the sample size. Therefore, our results have implications for efficient CVaR optimization.

(d) We show via numerical experiments with simulated and real data that the ARVaR formulation results in portfolio allocations that dominate the performance of other approaches for VaR optimization and sample CVaR optimization, and do not stretch the tail of the distribution exceeding VaR.

The structure of this paper is as follows: In §2, we review the main ideas behind the robust optimization methodology, and place the VaR optimization problem in this context. In §3, we discuss the asymmetric deviation framework of Chen et al. (2007), and show the formulation of the Asymmetry-Robust VaR Problem for a factor model of returns. In §4, we review the concept of coherent risk measures, and prove that the ARVaR formulation results in a coherent portfolio risk measure. In §5, we discuss the relationship between ARVaR and CVaR. Section 6 contains simulated and historical data computational experiments that compare the performance of the ARVaR formulation to the performance of other distribution-parameter-based and sample-based approaches for VaR optimization: worst-case VaR (El Ghaoui et al.

2003), VaR based on normal approximation, VaR resulting from sample CVaR minimization, and exact sample VaR.

Notation. We use boldface to denote vectors and matrices. Vectors are in lower case, whereas matrices are in upper case. Tilde ($\tilde{\cdot}$) denotes uncertain parameters.

2. Robust Optimization

Consider the following family of optimization problems:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & f(\mathbf{x}, \tilde{\mathbf{z}}) \geq 0, \\ & \mathbf{x} \in X. \end{aligned} \quad (3)$$

Here $\mathbf{x} \in \mathcal{R}^N$ is the vector of decision variables, $\tilde{\mathbf{z}} \in \mathcal{R}^M$ is a vector of uncertain factors, and the set X contains constraints whose parameters are certain. We can assume without loss of generality that the uncertain factors $\tilde{\mathbf{z}}$ satisfy the normalized distributional conditions $E(\tilde{\mathbf{z}}) = \mathbf{0}$ and $E(\tilde{\mathbf{z}}\tilde{\mathbf{z}}') = \mathbf{I}$. This can be achieved by a suitable linear transformation. For example, portfolio returns $\tilde{\mathbf{r}} \in \mathcal{R}^N$ with known means vector $\check{\mathbf{r}} \in \mathcal{R}^N$ and invertible covariance matrix $\Sigma \in \mathcal{R}^{N \times N}$ can be expressed as $\tilde{\mathbf{r}} = \check{\mathbf{r}} + \Sigma^{1/2}\tilde{\mathbf{z}}$ for some uncertain factors $\tilde{\mathbf{z}} \in \mathcal{R}^N$ satisfying the normalized distributional conditions. Hence, $\tilde{\mathbf{z}} = \Sigma^{-1/2}(\tilde{\mathbf{r}} - \check{\mathbf{r}})$. We will restrict our attention to robust linear constraints in which $f(\mathbf{x}, \mathbf{z})$ is bilinear in \mathbf{x} and $\tilde{\mathbf{z}}$; namely,

$$f(\mathbf{x}, \tilde{\mathbf{z}}) \triangleq f_0(\mathbf{x}) + \sum_{j=1}^M f_j(\mathbf{x})\tilde{z}_j, \quad (4)$$

where $f_j(\mathbf{x})$, $j = 0, \dots, M$ are linear functions in \mathbf{x} . For example, in the context of portfolio VaR optimization, the function in the constraint in formulation (2), $\tilde{\mathbf{r}}'\mathbf{x} + \gamma$, can be written in the form (4) in the following way:

$$\tilde{\mathbf{r}}'\mathbf{x} + \gamma = (\check{\mathbf{r}} + \Sigma^{1/2}\tilde{\mathbf{z}})'\mathbf{x} + \gamma = \underbrace{\check{\mathbf{r}}'\mathbf{x} + \gamma}_{f_0(\mathbf{x}, \gamma)} + \sum_{j=1}^M \underbrace{(\mathbf{e}_j'(\Sigma^{1/2})'\mathbf{x})}_{f_j(\mathbf{x}, \gamma)} \tilde{z}_j,$$

where $\mathbf{e}_j \in \mathcal{R}^M$ is a unit vector with one at the j th position and zeros in all other positions.

For any fixed solution \mathbf{x} , the constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \geq 0$ in optimization problem (3) may become infeasible for some realization of $\tilde{\mathbf{z}}$. The *robust optimization approach* to addressing this issue is to ensure that the solution \mathbf{x} satisfies the constraint for all realizations \mathbf{z} within a prespecified uncertainty set. This is accomplished by replacing the original constraint with its *robust counterpart*, which is defined as

$$f(\mathbf{x}, \mathbf{z}) \geq 0 \quad \forall \mathbf{z} \in \mathcal{U}(\tilde{\mathbf{z}}), \quad (5)$$

where $\mathcal{U}(\tilde{\mathbf{z}})$ is the uncertainty set that is mapped out from the probability distributions of the uncertain factors, $\tilde{\mathbf{z}}$.

The framework of scenario optimization can be incorporated in the robust counterpart by specifying $\mathcal{U}(\tilde{\mathbf{z}})$ as a collection of scenarios of $\tilde{\mathbf{z}}$. In that case, the robust counterpart is a set of constraints—one for each scenario for $\tilde{\mathbf{z}}$. However, the uncertainty set $\mathcal{U}(\tilde{\mathbf{z}})$ can be extended to richer sets, ranging from polytopes to more advanced conic-representable uncertainty sets. Such uncertainty sets can represent compactly an exponential or even an infinite number of scenarios. If the function $f(\mathbf{x}, \mathbf{z})$ is bilinear in (\mathbf{x}, \mathbf{z}) , which is the case for the portfolio optimization problem under consideration, convex uncertainty sets often lead to computationally tractable robust counterparts. The interested reader is referred to Ben-Tal and Nemirovski (1998) and El Ghaoui et al. (1998) for an overview of robust optimization problems.

The *size* of the uncertainty set is usually related to guarantees on the probability that the constraint involving uncertain coefficients will not be violated. Moreover, because in practice the exact distribution of the uncertain factors is rarely known, it is frequently convenient to specify a *shape* for the uncertainty set that reflects one's knowledge of the distribution, and that makes problem (5) efficiently solvable. Results on probability bounds related to the size and the shape of uncertainty sets can be found, for example, in Ben-Tal and Nemirovski (2001), Bertsimas and Sim (2004), Bertsimas et al. (2004), and Chen et al. (2007).

One of the most widely used kinds of uncertainty sets in robust optimization is the ellipsoidal set,

$$\mathcal{E}_\Omega(\tilde{\mathbf{z}}) = \{\mathbf{z}: \|\mathbf{z}\|_2 \leq \Omega\}. \quad (6)$$

For $\tilde{\mathbf{z}}$ in uncertainty set (6), the robust counterpart of a constraint of the form

$$f(\mathbf{x}, \tilde{\mathbf{z}}) \geq 0$$

can be derived using convex duality, and is

$$f_0(\mathbf{x}) - \Omega \cdot \sigma(f(\mathbf{x}, \tilde{\mathbf{z}})) \geq 0,$$

where $f_0(\mathbf{x})$ is the expected value and

$$\sigma(f(\mathbf{x}, \tilde{\mathbf{z}})) = \left\| \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_M(\mathbf{x}) \end{bmatrix} \right\|_2$$

is the standard deviation of $f(\mathbf{x}, \tilde{\mathbf{z}})$ for any given \mathbf{x} .

El Ghaoui et al. (2003) use this result, and specify an ellipsoidal uncertainty set for asset returns of a size determined by a parameter $\kappa = \sqrt{(1-\epsilon)/\epsilon}$ that guarantees that the VaR constraint in formulation (2) would be satisfied probabilistically for all possible

distributions for uncertain returns, and in particular, for the worst-case distribution with respect to probability measures with fixed first two moments, covariance matrix Σ , and vector of portfolio weights \mathbf{x} . They show equivalence between the worst-case $(1 - \epsilon)$ -VaR problem and the following formulation:

$$\begin{aligned}
 \text{(Worst-Case VaR)} \quad & \min \gamma \\
 \text{s.t.} \quad & \kappa \sqrt{\mathbf{x}' \Sigma \mathbf{x}} - \check{\mathbf{r}}' \mathbf{x} \leq \gamma, \\
 & \mathbf{x} \in \mathcal{X}. \tag{7}
 \end{aligned}$$

Here $\mathbf{x} \in \mathcal{X}$ includes additional constraints on the portfolio structure, such as the necessary finite budget constraint $\mathbf{x}' \mathbf{e} = 1$, and a constraint requiring a target expected future portfolio return, $\check{\mathbf{r}}' \mathbf{x} \geq r_{\text{target}}$.

If the uncertain factors $\tilde{\mathbf{z}}$ are assumed to be normally distributed, one can solve a less conservative version of model (7) in which $\kappa = \Phi^{-1}(1 - \epsilon)$ ($\Phi^{-1}(\cdot)$ stands for the inverse cumulative standard normal distribution). We will refer to the latter model as Normal VaR. This is a model that is often used in industry.

Both model (7) and the Normal VaR model rely on first- and second-moment information about the distribution of uncertainties. In the next section, we derive the robust counterpart of the portfolio VaR optimization problem (2) for an uncertainty set that relies on asymmetric measures of variability, and provides a tighter approximation to the actual VaR optimization solution.

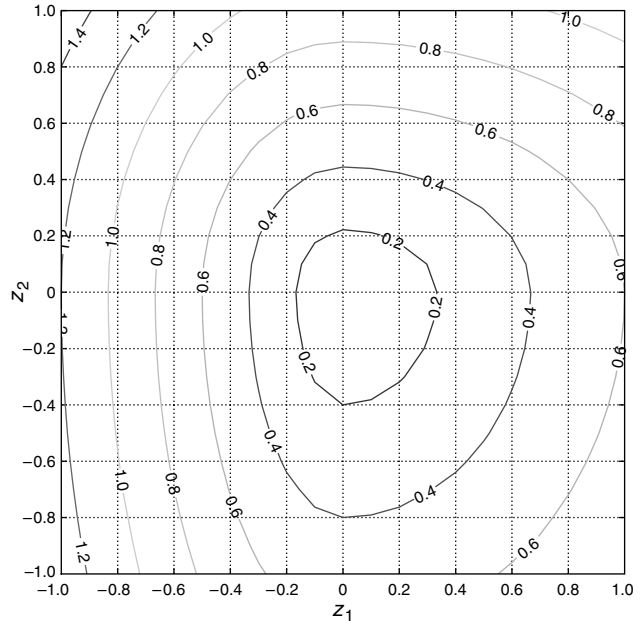
3. The Robust Framework of Chen et al. (2007)

Because the ellipsoidal uncertainty set \mathcal{E}_Ω is symmetric, it may be overly conservative for mapping asymmetric distributions for any particular level of required probabilistic guarantees. We consider instead the following uncertainty set, suggested by Chen et al. (2007) in the context of stochastic programming applications:

$$\begin{aligned}
 \mathcal{G}_\Omega(\tilde{\mathbf{z}}) = \{ \mathbf{z}: \exists \mathbf{v}, \mathbf{w} \in \mathcal{R}_+^M, \mathbf{z} = \mathbf{v} - \mathbf{w}, \\
 \|\mathbf{P}^{-1} \mathbf{v} + \mathbf{Q}^{-1} \mathbf{w}\| \leq \Omega, -\underline{\mathbf{z}} \leq \mathbf{z} \leq \bar{\mathbf{z}} \}, \tag{8}
 \end{aligned}$$

where $[-\underline{z}_j, \bar{z}_j]$ is the minimal set that contains \tilde{z}_j ($-\underline{z}_j$ and \bar{z}_j can be infinity), $\mathbf{P} = \text{diag}(p_1, \dots, p_M)$, and $\mathbf{Q} = \text{diag}(q_1, \dots, q_M)$. The parameters $p_j > 0$ and $q_j > 0$ are the “forward” and the “backward” deviations of random variable \tilde{z}_j , $j = 1, \dots, M$, respectively (Chen et al. 2007). The uncertainty set $\mathcal{G}_\Omega(\tilde{\mathbf{z}})$ is convex, and its size is controlled by Ω as illustrated in Figure 1. Intuitively speaking, the uncertain factors $\tilde{\mathbf{z}}$ are decomposed into two random variables: $\tilde{\mathbf{v}} = \max\{\tilde{\mathbf{z}}, 0\}$ and $\tilde{\mathbf{w}} = \max\{-\tilde{\mathbf{z}}, 0\}$, so that $\tilde{\mathbf{z}} = \tilde{\mathbf{v}} - \tilde{\mathbf{w}}$.

Figure 1 An Example of the Uncertainty Set \mathcal{G}_Ω as Ω Varies



The multipliers $1/p_j$ and $1/q_j$, $j = 1, \dots, M$, normalize the perturbation contributed by v_j or w_j , respectively, so that the norm of the total scaled deviation falls within the uncertainty budget Ω . As illustrated in Chen et al. (2007), if $\mathbf{P} = \mathbf{Q} = \mathbf{I}$ and $\underline{z}_j = \bar{z}_j = \infty$ for all $j = 1, \dots, M$, then $\mathcal{G}_\Omega(\tilde{\mathbf{z}}) = \mathcal{E}_\Omega(\tilde{\mathbf{z}})$. Hence, the asymmetric uncertainty set (8) generalizes the ellipsoidal uncertainty set (6) used by El Ghaoui et al. (2003) in their worst-case VaR formulation.

We discuss some useful properties of the asymmetric deviation measures, and show how to formulate an approximate VaR optimization problem for a factor model for returns based on uncertainty set (8).

3.1. Properties of the Asymmetric Deviation Measures

The forward and the backward deviation measures for a random variable \tilde{z} are defined as $p(\tilde{z}) = \inf\{\mathcal{P}(\tilde{z})\}$ and $q(\tilde{z}) = \inf\{\mathcal{Q}(\tilde{z})\}$, where (see Chen et al. 2007)

$$\mathcal{P}(\tilde{z}) = \left\{ \alpha: \alpha > 0, E\left(\exp\left(\frac{\phi}{\alpha} \tilde{z}\right)\right) \leq \exp\left(\frac{\phi^2}{2}\right) \forall \phi > 0 \right\} \tag{9}$$

and

$$\mathcal{Q}(\tilde{z}) = \left\{ \beta: \beta > 0, E\left(\exp\left(-\frac{\phi}{\beta} \tilde{z}\right)\right) \leq \exp\left(\frac{\phi^2}{2}\right) \forall \phi > 0 \right\}. \tag{10}$$

These deviation measures capture distributional asymmetry. It can be shown (Chen et al. 2007) that for a random variable \tilde{z} with zero mean, $p(\tilde{z})$ and $q(\tilde{z})$ are always greater than or equal to the standard deviation of the distribution. In general, $p(\tilde{z})$ and $q(\tilde{z})$ are finite if the support $[-\underline{z}, \bar{z}]$ of the distribution for \tilde{z}

is finite. If the support is infinite, $p(\tilde{z})$ and $q(\tilde{z})$ are not guaranteed to be finite. However, in the important case of a normally distributed random variable \tilde{z} , $p(\tilde{z})$ and $q(\tilde{z})$ are finite, and equal the standard deviation (Chen et al. 2007).

The exact distribution of \tilde{z} is unknown in many practical problems. It is therefore important to be able to estimate $p(\tilde{z})$ and $q(\tilde{z})$ from data. The following result provides a heuristic for doing so.¹

THEOREM 1. *Given a known distribution, or a set of data for the random variable \tilde{z} , the forward and backward deviations $p(\tilde{z})$ and $q(\tilde{z})$ can be determined as follows:*

$$p(\tilde{z}) = \sup_{\theta > 0} \left\{ \sqrt{2 \frac{\ln(\mathbb{E}(\exp(\theta\tilde{z})))}{\theta^2}} \right\} \quad (11)$$

and

$$q(\tilde{z}) = \sup_{\theta > 0} \left\{ \sqrt{2 \frac{\ln(\mathbb{E}(\exp(-\theta\tilde{z})))}{\theta^2}} \right\}. \quad (12)$$

PROOF. We show the proof related to the forward deviation. The result related to the backward deviation follows trivially. Observe that

$$\begin{aligned} \inf\{\mathcal{P}(\tilde{z})\} &= \inf \left\{ \alpha : \alpha > 0, \mathbb{E} \left(\frac{\phi}{\alpha} \tilde{z} \right) \leq \exp \left(\frac{\phi^2}{2} \right) \forall \phi > 0 \right\} \\ &= \inf \left\{ \alpha : \alpha > 0, \mathbb{E}(\exp(\rho\tilde{z})) \leq \exp \left(\frac{\rho^2 \alpha^2}{2} \right) \forall \rho > 0 \right\} \\ &\quad \text{(because } \rho = \phi/\alpha) \\ &= \inf \left\{ \alpha : \sqrt{2 \ln(\mathbb{E}(\exp(\rho\tilde{z})))}/\rho^2 \leq \alpha \forall \rho > 0 \right\} \\ &= \sup_{\theta > 0} \left\{ \sqrt{2 \frac{\ln(\mathbb{E}(\exp(\theta\tilde{z})))}{\theta^2}} \right\}. \quad \square \end{aligned}$$

The next two theorems are important results that will allow us to formulate the robust parametric portfolio VaR problem.

THEOREM 2. *The robust counterpart of the constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \geq 0$ when $\tilde{\mathbf{z}} \in \mathcal{R}^M$ is assumed to vary in the uncertainty set $\mathcal{G}_\Omega(\tilde{\mathbf{z}})$ and the distribution supports $[-\underline{z}_j, \bar{z}_j]$, $j = 1, \dots, M$, are finite, is equivalent to the following set of inequalities:*

$$\exists \mathbf{u}, \mathbf{r}, \mathbf{s} \in \mathcal{R}^M \begin{cases} f_0(\mathbf{x}) \geq \Omega \|\mathbf{u}\|_2 + \mathbf{r}'\tilde{\mathbf{z}} + \mathbf{s}'\underline{\mathbf{z}}, \\ u_j \geq -p_j(f_j(\mathbf{x}) + r_j - s_j), \quad u_j \geq q_j(f_j(\mathbf{x}) + r_j - s_j), \\ \quad j = 1, \dots, M, \\ \mathbf{r}, \mathbf{s} \geq \mathbf{0}. \end{cases} \quad (13)$$

¹The MATLAB code for implementing the heuristic is available from <http://www.bschool.nus.edu/STAFF/dscsim/research.htm>.

If the distribution support $[-\underline{z}_j, \bar{z}_j]$ is infinite, but p_j and q_j , $j = 1, \dots, M$, are finite, the first two sets of inequalities in (13) are replaced by

$$\begin{aligned} f_0(\mathbf{x}) &\geq \Omega \|\mathbf{u}\|_2, \\ u_j &\geq -p_j f_j(\mathbf{x}), \quad u_j \geq q_j f_j(\mathbf{x}), \quad j = 1, \dots, M. \end{aligned} \quad (14)$$

REMARK. In the interest of brevity, for the rest of the paper we will assume finite distribution support $[-\underline{z}_j, \bar{z}_j]$, $j = 1, \dots, M$. This is not a very restrictive assumption in practice, because the bounds can be made arbitrarily large. If the distributional support is unbounded (or the bounds are unknown), a modeler can replace the first two sets of inequalities in robust counterpart (13) with the inequalities in (14).

THEOREM 3. *Suppose that $\tilde{\mathbf{z}}$ are independently distributed, and \mathbf{x} is feasible in the robust counterpart (13). Then,*

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \geq 0) \geq 1 - \exp(-\Omega^2/2).$$

The proofs of Theorems 2 and 3 are contained in Chen et al. (2007).

3.2. The Asymmetry-Robust VaR Optimization Problem Formulation

Suppose that future asset returns $\tilde{\mathbf{r}} \in \mathcal{R}^N$ are generated by a factor model

$$\tilde{\mathbf{r}} = \check{\mathbf{r}} + \mathbf{A}\tilde{\mathbf{z}}, \quad (15)$$

in which $\check{\mathbf{r}} \in \mathcal{R}^N$ is a vector of expected returns, and $\mathbf{A} \in \mathcal{R}^{N \times M}$ is a matrix of factor loadings. The factors $\tilde{\mathbf{z}} \in \mathcal{R}^M$ have zero means and support $[-\underline{\mathbf{z}}, \bar{\mathbf{z}}]$, and are stochastically independent. We would like to find a vector of asset weights $\mathbf{x} \in \mathcal{R}^N$ that satisfies

$$\mathbb{P}(\gamma + \check{\mathbf{r}}'\mathbf{x} \geq 0) \geq 1 - \epsilon,$$

or equivalently,

$$\mathbb{P}(\gamma + \check{\mathbf{r}}'\mathbf{x} + \tilde{\mathbf{z}}'(\mathbf{A}'\mathbf{x}) \geq 0) \geq 1 - \epsilon.$$

Based on Theorems 2 and 3, if \mathbf{x} is feasible for the set of inequalities (13), then we can guarantee that

$$\mathbb{P}(\gamma + \check{\mathbf{r}}'\mathbf{x} \geq 0) \geq 1 - \exp(-\Omega^2/2). \quad (16)$$

Hence, a relaxation of the VaR optimization problem can be formulated as follows:

(Asymmetry-Robust VaR)

$$\begin{aligned} \min \quad & \gamma \\ \text{s.t.} \quad & \gamma + \check{\mathbf{r}}'\mathbf{x} \geq \Omega \|\mathbf{u}\|_2 + \mathbf{r}'\tilde{\mathbf{z}} + \mathbf{s}'\underline{\mathbf{z}}, \\ & \mathbf{y} = \mathbf{A}'\mathbf{x}; \quad \mathbf{r}, \mathbf{s} \geq \mathbf{0}, \\ & u_j \geq -p_j(y_j + r_j - s_j), \quad j = 1, \dots, M, \\ & u_j \geq q_j(y_j + r_j - s_j), \quad j = 1, \dots, M, \\ & \mathbf{x} \in \mathcal{X}. \end{aligned} \quad (17)$$

The optimal value γ^* computed from (17) (i.e., the ARVaR value) is a conservative approximation to the true portfolio VaR. Specifically, in view of (16), if a *minimum* confidence level of $(1 - \epsilon)$ for VaR is desired, Ω should be selected to be $\sqrt{-2 \ln \epsilon}$. For example, to compute the optimum 95% VaR for a portfolio, Ω should be set to 2.4477.

We note that (17) is a convex (second-order cone) optimization problem. Thus, the proposed approximation of VaR can be solved efficiently both in theory and in practice.

4. Coherent Risk Measures

The ARVaR formulation is attractive not only because of its computational tractability, but also from a financial theory perspective. As we will show in this section, ARVaR defines a coherent risk measure. This is an important result, because popular alternative approximation methods for VaR optimization do not have that property.

A risk measure $\rho(\cdot)$ is a functional defined on a random variable \tilde{v} . In the portfolio optimization context, the random variable is the return of a risky asset. By convention, $\rho(\tilde{v}) \leq 0$ implies that the risk associated with an uncertain return \tilde{v} is acceptable. Artzner et al. (1997) present and justify a set of four desirable properties for measures of risk, and call risk measure functionals $\rho(\cdot)$ that satisfy these properties *coherent*. The four desirable properties are as follows:

- (i) **Translation invariance:** For all $\tilde{v} \in \mathcal{V}$ and $a \in \mathcal{R}$, $\rho(\tilde{v} + a) = \rho(\tilde{v}) - a$.
- (ii) **Subadditivity:** For all random variables $\tilde{v}_1, \tilde{v}_2 \in \mathcal{V}$, $\rho(\tilde{v}_1 + \tilde{v}_2) \leq \rho(\tilde{v}_1) + \rho(\tilde{v}_2)$.
- (iii) **Positive homogeneity:** For all $\tilde{v} \in \mathcal{V}$ and $\lambda \geq 0$, $\rho(\lambda \tilde{v}) = \lambda \rho(\tilde{v})$.
- (iv) **Monotonicity:** For all random variables \tilde{v}_1, \tilde{v}_2 such that $\tilde{v}_1 \geq \tilde{v}_2$, $\rho(\tilde{v}_1) \leq \rho(\tilde{v}_2)$.

VaR is not a coherent risk measure by these rules because it violates the axiom of subadditivity. However, as we will demonstrate in this section, obtaining ARVaR as the optimal solution to formulation (17) guarantees a coherent portfolio risk measure.

Assuming factor model (15), the risky portfolio return \tilde{v} can be expressed as an affine function of the uncertain factors $\tilde{\mathbf{z}}$, that is, $\tilde{v} = v(\tilde{\mathbf{z}}) = f(\mathbf{x}, \tilde{\mathbf{z}})$, where $f(\mathbf{x}, \tilde{\mathbf{z}})$ was defined in Equation (4). We can thus define the following risk measure:

$$\eta_\Omega(\tilde{v}) \triangleq - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} v(\tilde{\mathbf{z}}) = - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} f(\mathbf{x}, \mathbf{z}). \quad (18)$$

Observe that

$$\begin{aligned} \eta_\Omega(\tilde{v}) \leq 0 &\Leftrightarrow - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} f(\mathbf{x}, \mathbf{z}) \leq 0 \\ &\Leftrightarrow \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} f(\mathbf{x}, \mathbf{z}) \geq 0 \\ &\Leftrightarrow f(\mathbf{x}, \mathbf{z}) \geq 0 \quad \forall \mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}}). \end{aligned}$$

Therefore,

$$\eta_\Omega(f(\mathbf{x}, \tilde{\mathbf{z}})) \leq 0$$

is equivalent to the robust counterpart (13).

THEOREM 4. *Under factor model (15) and assuming that the factors $\tilde{z}_j, j = 1, \dots, M$, are independently distributed, the risk measure $\eta_\Omega(\cdot)$ is a coherent risk measure defined on the portfolio returns.*

PROOF. To show translation invariance, we observe that

$$\begin{aligned} \eta_\Omega(v(\tilde{\mathbf{z}}) + a) &= - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} (v(\mathbf{z}) + a) \\ &= -a - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} v(\mathbf{z}) \\ &= \eta_\Omega(v(\tilde{\mathbf{z}})) - a. \end{aligned}$$

To show positive homogeneity, we note that for all $\lambda \geq 0$,

$$\begin{aligned} \eta_\Omega(\lambda v(\tilde{\mathbf{z}})) &= - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} (\lambda v(\mathbf{z})) \\ &= -\lambda \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} v(\mathbf{z}) \\ &= \lambda \eta_\Omega(v(\tilde{\mathbf{z}})). \end{aligned}$$

To show compliance with the subadditivity axiom, we consider the random returns on two portfolios, $v(\tilde{\mathbf{z}})$ and $w(\tilde{\mathbf{z}})$:

$$\begin{aligned} \eta_\Omega(v(\tilde{\mathbf{z}}) + w(\tilde{\mathbf{z}})) &= - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} (v(\mathbf{z}) + w(\mathbf{z})) \\ &\leq - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} v(\mathbf{z}) - \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} w(\mathbf{z}) \\ &= \eta_\Omega(v(\tilde{\mathbf{z}})) + \eta_\Omega(w(\tilde{\mathbf{z}})). \end{aligned}$$

To show that the axiom of monotonicity is satisfied, we will first show that if the portfolio return is always nonnegative, that is, $v(\tilde{\mathbf{z}}) \geq 0$, then $\eta_\Omega(v(\tilde{\mathbf{z}})) \leq 0$. Because the factors \tilde{z}_j are independently distributed and $[-\underline{z}_j, \bar{z}_j]$ is the minimal set that contains \tilde{z}_j , the minimal convex set containing $\tilde{\mathbf{z}}$ is $\mathcal{W}(\tilde{\mathbf{z}}) = [-\underline{\mathbf{z}}, \bar{\mathbf{z}}]$. Given a portfolio return $v(\tilde{\mathbf{z}})$ that satisfies $v(\tilde{\mathbf{z}}) \geq 0$, we will argue that

$$v(\mathbf{z}) \geq 0 \quad \forall \mathbf{z} \in \mathcal{W}(\tilde{\mathbf{z}}). \quad (19)$$

Suppose not, i.e., suppose that there exists a $\mathbf{y} \in \mathcal{W}(\tilde{\mathbf{z}})$ such that $v(\mathbf{y}) < 0$. Because \mathbf{y} is in the convex hull of the space of possible realizations of $\tilde{\mathbf{z}}$, there exist positive λ_k satisfying $\sum_{k=1}^K \lambda_k = 1$ and realizations \mathbf{z}^k such that

$$\mathbf{y} = \sum_{k=1}^K \lambda_k \mathbf{z}^k.$$

Under factor model (15), the portfolio return $v(\tilde{\mathbf{z}})$ is an affine function of $\tilde{\mathbf{z}}$. Therefore,

$$v(\mathbf{y}) = \sum_{k=1}^K \lambda_k v(\mathbf{z}^k) < 0,$$

which is a contradiction because $v(\tilde{\mathbf{z}}) \geq 0$. Because $\mathcal{G}_\Omega(\tilde{\mathbf{z}}) \subseteq \mathcal{W}(\tilde{\mathbf{z}})$ for all $\Omega > 0$, inequality (19) implies

$$v(\mathbf{z}) \geq 0 \quad \forall \mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}}),$$

or equivalently, $\eta_\Omega(v(\tilde{\mathbf{z}})) \leq 0$. Finally, given two portfolio returns $v(\tilde{\mathbf{z}})$ and $w(\tilde{\mathbf{z}})$, such that $v(\tilde{\mathbf{z}}) \geq w(\tilde{\mathbf{z}})$, we have

$$\begin{aligned} \eta_\Omega(v(\tilde{\mathbf{z}})) &= \eta_\Omega(v(\tilde{\mathbf{z}}) - w(\tilde{\mathbf{z}}) + w(\tilde{\mathbf{z}})) \\ &\leq \eta_\Omega(\underbrace{v(\tilde{\mathbf{z}}) - w(\tilde{\mathbf{z}})}_{\geq 0}) + \eta_\Omega(w(\tilde{\mathbf{z}})) \quad (\text{subadditivity}) \\ &\leq \eta_\Omega(w(\tilde{\mathbf{z}})), \end{aligned}$$

i.e., the monotonicity axiom is satisfied. \square

REMARK. While we proved Theorem 4 by considering the worst-case return over a deterministic uncertainty set, our results can be related also to a large body of literature on the relationship between coherent risk measures and worst-case expected portfolio return over a family of probability distributions.² The so-called *representation theorem* (see, for example, Föllmer and Schied 2002) states that a risk measure $\rho(\cdot)$ is coherent if and only if there exists a family of probability measures \mathcal{Q} such that

$$\rho(\tilde{v}) = \sup_{Q \in \mathcal{Q}} E_Q(-\tilde{v}),$$

where $E_Q(\tilde{v})$ denotes the expectation of the random variable \tilde{v} under the measure Q (as opposed to the measure of \tilde{v} itself).

In our context, \tilde{v} is the portfolio return and $\tilde{v} = f(\mathbf{x}, \tilde{\mathbf{z}})$, where $f(\mathbf{x}, \tilde{\mathbf{z}}) = \tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'\mathbf{A}\tilde{\mathbf{z}}$ from (4) and (15). Next, we show how it can be established that $\eta_\Omega(\cdot)$ is a coherent risk measure according to the representation theorem.

Let Λ be the sample space of random factors, and \mathcal{P} be the set of all possible probability measures defined on the sample space Λ . Based on the definition of $\eta_\Omega(\cdot)$ in (18) and the fact that

$$-\min_{\mathbf{z} \in G_\Omega(\tilde{\mathbf{z}})} f(\mathbf{x}, \tilde{\mathbf{z}}) = \max_{\mathbf{z} \in G_\Omega(\tilde{\mathbf{z}})} -(\tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'\mathbf{A}\tilde{\mathbf{z}}),$$

it suffices to show that

$$\max_{\mathbf{z} \in G_\Omega(\tilde{\mathbf{z}})} -\tilde{\mathbf{z}}'(\mathbf{A}'\mathbf{x}) = \max_{Q \in \mathcal{Q}} E_Q(-\tilde{\mathbf{z}}'(\mathbf{A}'\mathbf{x}))$$

for some family of distributions $\mathcal{Q} \subseteq \mathcal{P}$. Observe that because \tilde{z}_j are independently distributed in $[-\underline{z}_j, \bar{z}_j]$, the convex hull of Λ is given by $\mathcal{W} = [-\underline{\mathbf{z}}, \bar{\mathbf{z}}]$. Hence, for any $\mathbf{z} \in \mathcal{W}$ that may not be in the sample space Λ , there exists a probability measure $Q \in \mathcal{P}$ such that

$E_Q(\tilde{\mathbf{z}}) = \mathbf{z}$. For all $\mathbf{z} \in \mathcal{W}$, let $D(\mathbf{z})$ be a probability measure in \mathcal{P} such that

$$E_{D(\mathbf{z})}(\tilde{\mathbf{z}}) = \mathbf{z}.$$

Because $G_\Omega(\tilde{\mathbf{z}}) \subseteq \mathcal{W}$, we can define a family of probability measures

$$\mathcal{Q} = \{Q \in \mathcal{P}: Q = D(\mathbf{z}) \text{ for some } \mathbf{z} \in G_\Omega(\tilde{\mathbf{z}})\}.$$

This ensures that, indeed,

$$\begin{aligned} \sup_{Q \in \mathcal{Q}} E_Q(-\tilde{\mathbf{z}}'(\mathbf{A}'\mathbf{x})) &= \sup_{Q \in \mathcal{Q}} -E_Q(\tilde{\mathbf{z}}'(\mathbf{A}'\mathbf{x})) \\ &= \max_{\mathbf{z} \in G_\Omega(\tilde{\mathbf{z}})} -\mathbf{z}'(\mathbf{A}'\mathbf{x}). \quad \blacksquare \end{aligned}$$

Observe that the problem of minimizing $\eta_\Omega(f(\mathbf{x}, \tilde{\mathbf{z}}))$ over $\mathbf{x} \in \mathcal{X}$, where $f(\mathbf{x}, \tilde{\mathbf{z}}) = \tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'(\mathbf{A}'\mathbf{x})$ from Equation (4), can be stated as

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \{ \gamma: \eta_\Omega(f(\mathbf{x}, \tilde{\mathbf{z}})) \leq \gamma \} \\ &= \min_{\mathbf{x} \in \mathcal{X}} \{ \gamma: \eta_\Omega(f(\mathbf{x}, \tilde{\mathbf{z}}) + \gamma) \leq 0 \} \\ &\quad (\text{translation invariance}) \\ &= \min_{\mathbf{x} \in \mathcal{X}} \{ \gamma: \eta_\Omega(\tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'(\mathbf{A}'\mathbf{x}) + \gamma) \leq 0 \} \\ &= \min_{\mathbf{x} \in \mathcal{X}} \left\{ \gamma: -\min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} \{ \tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'(\mathbf{A}'\mathbf{x}) + \gamma \} \leq 0 \right\} \\ &= \min_{\mathbf{x} \in \mathcal{X}} \left\{ \gamma: \min_{\mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}})} \{ \tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'(\mathbf{A}'\mathbf{x}) + \gamma \} \geq 0 \right\} \\ &= \min_{\mathbf{x} \in \mathcal{X}} \{ \gamma: \tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'(\mathbf{A}'\mathbf{x}) + \gamma \geq 0 \quad \forall \mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}}) \} \\ &= \min_{\mathbf{x} \in \mathcal{X}} \{ \gamma: \tilde{\mathbf{r}}'\mathbf{x} + \mathbf{z}'\mathbf{y} + \gamma \geq 0 \quad \forall \mathbf{z} \in \mathcal{G}_\Omega(\tilde{\mathbf{z}}), \mathbf{y} = \mathbf{A}'\mathbf{x} \}. \end{aligned}$$

Applying Theorem 2, we obtain formulation (17). In other words, by optimizing (17), we obtain a coherent approximation to portfolio VaR.

Note that the ellipsoidal uncertainty set $\mathcal{E}_\Omega(\tilde{\mathbf{z}})$ does not necessarily satisfy $\mathcal{E}_\Omega(\tilde{\mathbf{z}}) \subseteq \mathcal{W}(\tilde{\mathbf{z}})$ for all $\Omega > 0$, unless the support $[-\underline{\mathbf{z}}, \bar{\mathbf{z}}]$ is infinite. Hence, risk measures associated with first and second moments such as the Worst-Case VaR (El Ghaoui et al. 2003) or the Normal VaR do not provide coherent approximations for VaR optimization because they violate the monotonicity axiom for a class of realistic probability distributions for returns.

5. The Relationship Between ARVaR and CVaR

CVaR is an example of a quantile-based risk measure that is also coherent. It measures the expected loss that can happen if the portfolio loss ends up higher than the $(1 - \epsilon)$ -quantile of the portfolio loss

² The authors thank the associate editor for pointing out this fact.

distribution. Mathematically, the $(1 - \epsilon)$ -CVaR measure is defined as (Rockafellar and Uryasev 2000)

$$\phi_{1-\epsilon}(\tilde{v}) \triangleq \min_a \left(a + \frac{E(-\tilde{v} - a)^+}{\epsilon} \right). \quad (20)$$

Under the CVaR measure, the portfolio optimization model becomes

$$\begin{aligned} \text{(CVaR)} \quad & \min \gamma \\ & \text{s.t. } \phi_{1-\epsilon}(\tilde{\mathbf{r}}'\mathbf{x} + \gamma) \leq 0, \\ & \mathbf{x} \in \mathcal{X}. \end{aligned} \quad (21)$$

Given the connection with the probabilistic constraint in terms of the parameter of reliability, $(1 - \epsilon)$, the $(1 - \epsilon)$ -CVaR measure can be viewed as an approximation of the chance constraint. In other words,

$$\phi_{1-\epsilon}(\tilde{v}) \leq 0 \Rightarrow P(\tilde{v} \geq 0) \geq 1 - \epsilon.$$

Unfortunately, exact evaluation of $\phi_{1-\epsilon}(\tilde{v})$ even for a simple linear function \tilde{v} requires multidimensional integration, which is computationally expensive. The problem of whether an approximation can be done efficiently is still open. However, for discrete distributions, the formulation is substantially simpler (Rockafellar and Uryasev 2000). Thus, the formulation of the portfolio CVaR problem that has been most useful in practice is the sample CVaR formulation. Given T realizations of returns $\{\mathbf{r}^1, \dots, \mathbf{r}^T\}$, the problem of finding a portfolio allocation that results in the minimum CVaR is a linear optimization problem, in which the constraint $\phi_{1-\epsilon}(\tilde{\mathbf{r}}'\mathbf{x} + \gamma) \leq 0$ is replaced by the set of constraints

$$\begin{aligned} -a + \frac{1}{K} \sum_{t=1}^T y_t &\leq \gamma, \\ a - y_t &\leq \mathbf{r}_t'\mathbf{x}, \quad t = 1, \dots, T, \\ y_t &\geq 0, \quad t = 1, \dots, T. \end{aligned} \quad (22)$$

The number of observations needed to achieve any meaningful confidence of reliability is of the order of $1/\epsilon$. In practice, the choice of the reliability parameter is often arbitrary, e.g., $\epsilon = 1\%$ or 5% . Therefore, the computation of CVaR may not scale well with portfolio variability. This observation agrees with empirical findings in Yamai and Yoshida (2002) that estimation of CVaR can be very unstable for relatively small samples, and is explored further in the computational experiments in §6. Incidentally, ARVaR gives an approximation to CVaR:

THEOREM 5 (CHEN AND SIM 2006). *Suppose that $\tilde{\mathbf{z}}$ is a vector of independently distributed factors, and \mathbf{x} is feasible in the set of inequalities (13). Then,*

$$\phi_{1-\epsilon}(f(\mathbf{x}, \tilde{\mathbf{z}})) \leq 0,$$

where $\epsilon = \exp(-\Omega^2/2)$.

The proof can be found in Chen and Sim (2006) and is quite involved, so we omit it here. The connection between ARVaR and CVaR creates new opportunities for efficient CVaR minimization.

6. Computational Experiments

We test the viability of the proposed parametric approach to VaR optimization using simulated and real market data. The controlled experiments with simulated data are designed to study the ARVaR performance when one has perfect information about the distribution of returns, and can estimate the parameters for all parametric VaR optimization approaches exactly. The market returns data experiments investigate whether ARVaR is useful in real-world situations with imperfect information. In both sets of experiments, we compare the performance of the ARVaR formulation to the following alternative approaches for VaR optimization:

- (i) Worst-case VaR (WVaR), computed by optimizing (7);
- (ii) Normal VaR (NVaR), which tends to be the standard in the financial industry;
- (iii) CVaR-based VaR, computed as the realized VaR given portfolio weights resulting from minimization of CVaR for a sample of returns (formulation (21) with constraints (22)). There is evidence that minimization of CVaR leads to minimization of VaR as well (Rockafellar and Uryasev 2002). In addition, several heuristics for computing VaR in the literature are based on approximating VaR by CVaR (see, for example, Larsen et al. 2002);
- (iv) Exact sample portfolio VaR (ESVaR), which can be computed by solving the following mixed-integer (MIP) optimization problem:

$$\begin{aligned} \min \quad & \gamma \\ \text{s.t. } & \gamma + (\mathbf{r}^i)'\mathbf{x} \geq -Ky_i, \quad i = 1, \dots, T, \\ & \mathbf{x}'\mathbf{e} = 1, \\ & \mathbf{y}'\mathbf{e} = \lfloor \epsilon T \rfloor, \\ & \mathbf{y} \in \{0, 1\}^T \end{aligned}$$

for some large constant K and a sample of T vectors of realized asset returns. To see that the optimal γ in the above problem is indeed the optimal $(1 - \epsilon)$ -VaR, note that in the optimal solution, $y_i = 1$ for the highest $\lfloor \epsilon T \rfloor$ realizations of portfolio losses (negative returns), and $y_i = 0$ for the remaining realizations of portfolio losses. For the set of indices i for which $y_i = 0$, the constraints

$$\gamma \geq -(\mathbf{r}^i)'\mathbf{x}$$

ensure that the optimal γ will take the lowest value that is higher than the lowest $(1 - \epsilon)T$ of all losses, which is exactly the definition of sample $(1 - \epsilon)$ -VaR.

For the remaining indices i (those for which $y_i = 1$), the constraints

$$\gamma + (\mathbf{r}^i)' \mathbf{x} \geq -Ky_i$$

are not binding.

As we mentioned earlier, VaR optimization is not a convex problem, and its sample approximation is itself a very intractable problem, so we set a time limit of 1,800 seconds to the solver (CPLEX) we used to solve the above problem for a given sample of data. The solution we obtain, albeit not guaranteed to be the optimal, is still useful for comparison purposes.

6.1. Controlled Experiments with Simulated Data

We consider a portfolio of $N = 24$ assets with uncertain returns \tilde{r}_i , $i = 1, \dots, N$. Each return \tilde{r}_i is determined by a simple single factor model $\tilde{r}_i = \check{r}_i + \tilde{z}_i$, where $\check{r}_i = 1$. The factors \tilde{z}_i are independent and distributed as follows:

$$\tilde{z}_i = \begin{cases} \frac{\sqrt{\beta_i(1-\beta_i)}}{\beta_i} & \text{with probability } \beta_i, \\ -\frac{\sqrt{\beta_i(1-\beta_i)}}{(1-\beta_i)} & \text{with probability } (1-\beta_i). \end{cases}$$

Note that all random stock returns \tilde{r}_i have the same mean and standard deviation. However, depending on the parameters β_i , $i = 1, \dots, N$, the degree of symmetry of each individual return distribution can be different. Higher values for β_i (e.g., $\beta_i = 0.9$) result in large losses and small upside gains. We generate values for β_i as follows:

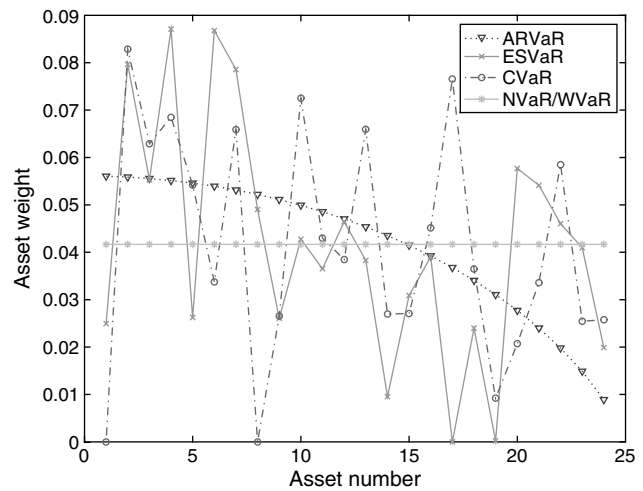
$$\beta_i = \frac{1}{2} \left(1 + \frac{i}{N+1} \right), \quad i = 1, \dots, N.$$

Therefore, the return distributions for stocks with high index numbers in the portfolio are more negatively skewed than those for stocks with low index numbers.

Despite its simplicity, this example is relevant for practical applications. One can imagine, for example, modeling credit defaults or credit ratings migration, where the return on a risky asset is dramatically influenced by the occurrence of a single discrete event.

We use exact (as opposed to estimated) values for the parameters in the ARVaR, WVaR, and NVaR optimization problems. These parameters include the covariance matrices and average returns for the WVaR and the NVaR approaches, and the bounds on the factor realizations as well as the backward and forward deviations for the ARVaR approach. We use a *training set* of 1,000 simulated returns from the above distributions as an input to the sample-based CVaR and ESVaR approaches. The optimal portfolio allocations resulting from the five approximate VaR

Figure 2 Optimal Portfolio Weights (as Proportions) for Assets Numbered 1 Through 24 Resulting from the Five Approximate VaR Optimization Methods in the Simulated Experiments ($\epsilon = 1\%$)



optimization approaches for $\epsilon = 1\%$ are shown in Figure 2. Because the asset standard deviations and means are the same, WVaR and NVaR do not distinguish between the assets, and allocate the wealth equally among them. ARVaR is able to detect the asymmetry in the distributions, and allocates less in assets with more negatively skewed return distributions (those with high index numbers). The behavior of the sample-based CVaR and ESVaR approaches is erratic. In fact, the optimal weights for the portfolios found with the latter two approaches vary widely from sample to sample. Figure 3 illustrates the variability in the optimal weights obtained from the sample CVaR optimization problem if 1,000 different sets of 1,000 scenarios each are generated.

Figure 3 Ranges of Variability of the Optimal Portfolio Weights (as Proportions) for Assets Numbered 1 Through 24 Resulting from Optimizing Sample CVaR on 1,000 Different Sets of 1,000 Independent Scenarios Each ($\epsilon = 1\%$)

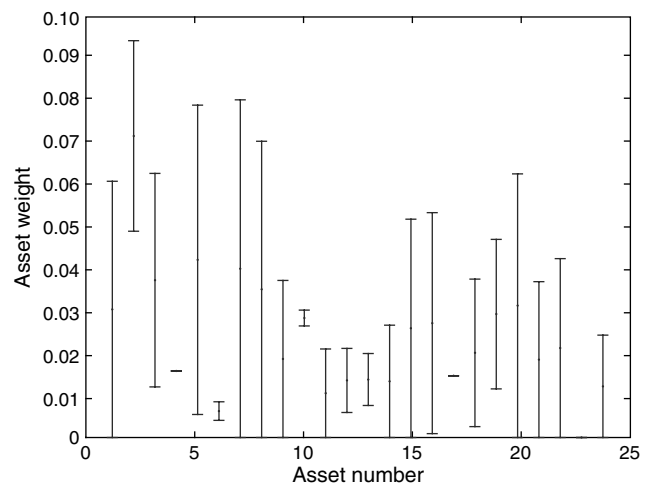


Table 1 Realized Portfolio VaR for the In-Sample and Out-of-Sample Simulated Experiments

ϵ (%)	In-sample realized VaR					Out-of-sample realized VaR				
	ARVaR	ESVaR	CVaR	NVaR	WVaR	ARVaR	ESVaR	CVaR	NVaR	WVaR
5	-0.420	-1.143	-0.558	-0.664	-0.665	-0.658	-1.143	-0.621	-0.646	-0.645
1	-0.281	-0.578	-0.481	-0.525	-0.524	-0.468	-0.414	-0.414	-0.477	-0.479
0.1	-0.120	-0.504	-0.514	-0.369	-0.369	-0.286	-0.208	-0.201	-0.272	-0.273
0.01	0.017	-0.527	-0.434	-0.241	-0.243	-0.139	-0.069	0.051	-0.109	-0.102

Notes. There are 1,000 observations in the training set, and 500,000 observations in the test set. The parameters for the ARVaR, NVaR, and WVaR optimization problems are obtained using perfect information about the distributions of uncertain returns. The sample CVaR and ESVaR optimization formulations are solved using the sample observations in the training set.

In Table 1, we list the realized in-sample VaR values for the original training set of 1,000 observations, as well as the realized out-of-sample VaR values for a *test set* of 500,000 new realizations of return vectors. We note that the numbers in the computational output for VaR represent the worst portfolio loss per dollar invested that can happen with probability ϵ , and it is therefore desirable to have low values for the VaR (negative VaR values mean that the portfolio loss at the $(1 - \epsilon)$ percentile is a gain, not a loss). To compute the in-sample realized VaR, we used the optimal weights found with the optimization formulations for the five VaR approximation methods, and computed the realized portfolio return for each observation in the training set. We then computed the $(1 - \epsilon)$ -percentile of portfolio losses (we assumed that the portfolio started out with 0% in each asset). To compute the out-of-sample realized VaR, we used the same optimal weights, but computed the portfolio returns for the return realizations in the test set, and took the $(1 - \epsilon)$ -percentile of portfolio losses for that distribution.

One can observe that while the sample-based approaches (CVaR and ESVaR) perform well in-sample, they lose ground to the distribution-parameter-based approaches (ARVaR, NVaR, and WVaR) in the out-of-sample experiments. Interestingly, the relative performance of the sample-based approaches is worse for small than for large values of ϵ . Small values of ϵ leave a limited number of sampled observations in the tail of the distribution, and thus the estimate of the expected losses in the tail is less reliable.

6.2. Experiments with Real Market Data

We consider a portfolio of 24 small cap stocks from different industry categories of the S&P 600 index (Table 2), and use historical returns from April 17, 1998 to June 1, 2006.³ Small cap stocks behave more erratically than large cap stocks, and tend to have more skewed historical return distributions. In fact,

³ The exact data set we used, as well as the descriptive statistics, are posted at <http://www.bschool.nus.edu/STAFF/dscsim/research.htm>.

all stocks in our sample have historical return distributions that are quite far from normal, and the factors we extract from the data are also asymmetrically distributed.

There are a total of 2,034 observations for each stock. We use the first 80% of these observations as a training set: they represent the sample returns in the CVaR and the ESVaR formulations, and are used to estimate the parameters for all optimization models. For the ARVaR formulation, we assume that returns are given by factor model (15), and that $\mathbf{A} = \mathbf{\Sigma}^{1/2}$, where $\mathbf{\Sigma}$ is the covariance matrix of the sample returns in the training set. We then compute uncorrelated factors $\tilde{\mathbf{z}} = \mathbf{\Sigma}^{-(1/2)}(\tilde{\mathbf{r}} - \tilde{\mathbf{r}})$, and assume they are also stochastically independent. While independence is not necessary for coherence of the ARVaR risk measure, it is required for the probability bound on the VaR constraint violation (16) used in §3.2. However, the computational results suggest that the probability bound is not a concern. The creation of a factor model itself is a challenging task—in practice, portfolio managers could use more sophisticated proprietary factor models for stock returns.

We use the last 20% of the data as a test set, and conduct in-sample and out-of-sample experiments similarly to the way we did in §6.1. Because the in-sample and out-of-sample results are consistent for different

Table 2 List of Stocks and Corresponding Industries Used in the Computational Experiments

Industry	Company name (ticker)
Consumer discretionary	Aztar Corp. (AZR), Finish Line (FNL), Hancock Fabrics Inc. (HKF), Haverly Furniture (HVT), K2 Inc. (KTO), Landry's Restaurants, Inc. (LNY)
Financials	Downey S & L Assn. (DSL), Harbor Florida Bancshares (HARB), Wintrust Financial Corp. (WTFC)
Industrials	AAR Corp. (AIR), CDI Corp. (CDI), ElkCorp (ELK), Frontier Airlines Holdings (FRNT), Gencorp (GY), Harland (J.H.) (JH), Healthcare Services Group (HCSG)
Information technology	FEI Company (FEIC), Exar Corp. (EXAR), Gevity HR (GVHR)
Healthcare	BioLase Technology (BLTI), Bradley Pharmaceuticals (BDR), Diagnostic Products (DP), Immucor Inc. (BLUD)
Materials	Penford Corp. (PENX)

Table 3 Forecasted and Realized VaR Values for $\epsilon = 1\%$ in the In-Sample Experiments

Target return (%)	WVaR	Real WVaR	NVaR	Real NVaR	ARVaR	Real ARVaR	CVaR-based VaR	ESVaR	Real ESVaR
0.300	27.116	5.868	6.110	5.868	9.660	3.591	3.447	2.835	3.182
0.500	49.933	11.130	11.292	11.129	14.836	5.490	5.195	4.095	4.580
0.600	61.638	13.778	13.952	13.778	17.424	6.748	6.064	4.764	5.707
0.800	85.235	19.243	19.316	19.249	22.600	9.556	8.140	6.363	7.241
1.000	108.946	24.813	24.706	24.810	29.759	11.768	10.127	8.059	9.977
1.200	132.710	30.276	30.109	30.275	37.358	13.707	12.064	9.600	11.324
1.400	156.503	35.739	35.519	35.738	44.957	15.688	14.325	11.255	13.034
1.600	180.314	41.202	40.933	41.202	52.556	17.669	16.586	12.972	14.824
2.000	227.965	52.128	51.767	52.129	67.755	22.379	20.848	16.433	18.716
2.500	287.561	65.786	65.318	65.786	86.752	29.135	26.191	20.747	24.330
3.000	347.174	79.248	78.873	79.257	105.750	34.976	31.676	25.480	32.165

Note. The realized values are denoted by “Real.” In the case of the CVaR-based VaR, only the realized VaR is reported.

values of ϵ (we experimented with $\epsilon = 0.1\%$, 1% , 5% , and 10%), in the interest of space we report them only for the case of $\epsilon = 1\%$. Table 3 contains the forecasted and the realized values of the portfolio VaR for the in-sample experiments. The forecasted values are the optimal objective function values in the different optimization formulations—they stand for the VaR values a portfolio manager expects to obtain. Figure 4 illustrates the in-sample realized portfolio efficient frontiers for the different optimization formulations. Similarly, Table 4 contains the realized value of the portfolio VaR for the out-of-sample experiments, and Figure 5 shows the out-of-sample realized portfolio efficient frontiers.

It is apparent that ARVaR outperforms both WVaR and NVaR in terms of realized VaR (WVaR and NVaR result in the same optimal portfolio allocation in our experiments because the constraint on a minimum target expected return is binding at the optimal solution, so both the WVaR and the NVaR formulations

reduce to minimizing the portfolio standard deviation). The dominance of ARVaR over WVaR and NVaR is particularly strong in the out-of-sample experiments. It is also interesting to observe that all realized VaR values for the ARVaR method are lower than the forecasted values from the optimization problem (this is not always the case, for example, for NVaR, especially for high values of the target return). The realized VaR values from the ARVaR method also are in the neighborhood of the ESVaR values, which cannot be said for the WVaR or NVaR realized values. This is encouraging, because it suggests that the approximation in the ARVaR formulation is tight.

As expected, because CVaR works directly with the data, it performs very well relative to all the other methods (with the exception of ESVaR) in the in-sample experiments. However, the danger of overfitting the data when it comes to sample-based portfolio optimization approaches is from the CVaR performance in the out-of-sample experiments—in those experiments, CVaR is dominated by ARVaR.

We mentioned earlier that there is evidence that optimizing VaR may “stretch” the tail of the portfolio loss distribution, resulting in very extreme maximum losses. We compare the maximum out-of-sample

Figure 4 In-Sample Realized Portfolio VaR Efficient Frontiers for the Different Optimization Formulations ($\epsilon = 1\%$)

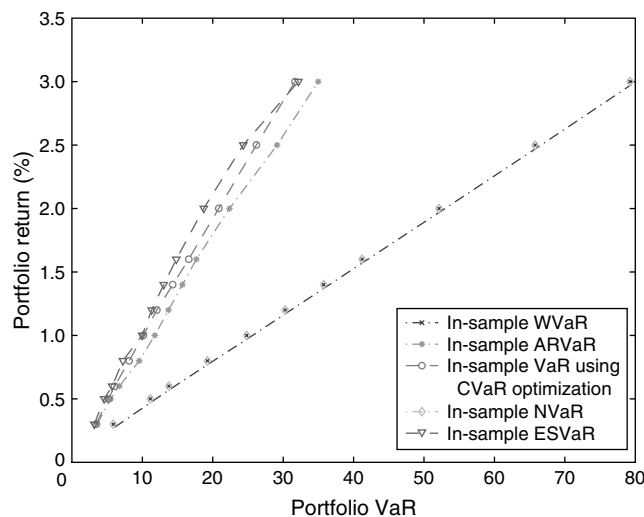
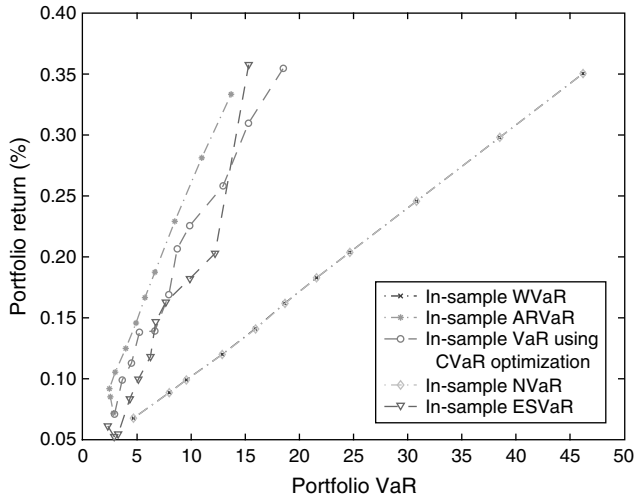


Table 4 Realized VaR Values for $\epsilon = 1\%$ in the Out-of-Sample Experiments

Target return (%)	Real WVaR	Real NVaR	Real ARVaR	CVaR-based VaR	Real ESVaR
0.300	4.661	4.661	2.810	2.909	2.292
0.500	7.971	7.968	2.527	3.622	2.878
0.600	9.546	9.547	2.444	4.479	3.251
0.800	12.863	12.865	2.978	5.211	4.327
1.000	15.912	15.953	3.950	6.640	5.117
1.200	18.644	18.668	4.905	7.934	6.232
1.400	21.586	21.583	5.727	8.705	6.713
1.600	24.663	24.658	6.645	9.867	7.646
2.000	30.818	30.815	8.483	12.946	9.882
2.500	38.510	38.513	10.972	15.296	12.203
3.000	46.205	46.208	13.691	18.509	15.312

Figure 5 Out-of-Sample Realized Portfolio VaR Efficient Frontiers for the Different Optimization Formulations ($\epsilon = 1\%$)



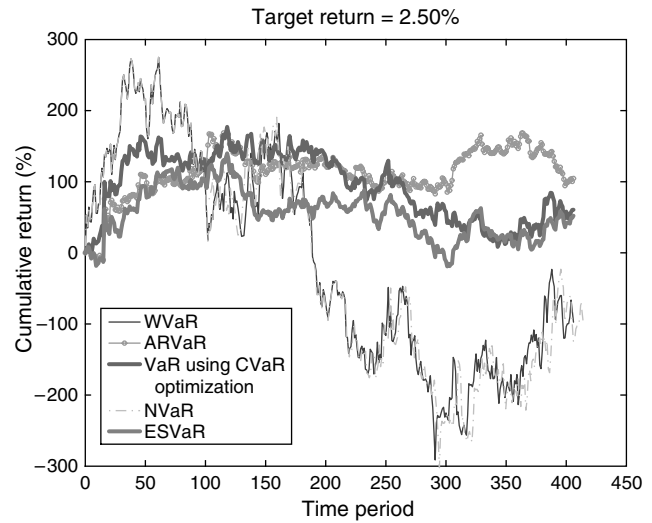
realized portfolio losses for the five VaR optimization approaches in Table 5. ARVaR results in the lowest out-of-sample maximum realized losses.

Finally, we study the cumulative portfolio return if a portfolio manager employs a simple buy-and-hold strategy. We assume that he allocates the portfolio assets according to the optimal weights from the five different portfolio formulations using parameters found from the first 80% of the data, and keeps the same weights for the remaining 20% of the time periods. In other words, he invests on October 21, 2004 (the 1,629th observation in our data set, and the first observation in the test set), and holds the same portfolio until June 1, 2006 (the last observation in the test set). The portfolio cumulative return graph for a target return of 2.5% is shown in Figure 6 (the graphs for other target percentage returns look very similar). The optimal portfolio allocation based on the ARVaR approach tends to result in very stable returns, whereas, for example, the behavior of the optimal portfolio obtained with the NVaR or the WVaR approaches is very erratic.

Table 5 Maximum Realized Out-of-Sample Portfolio Losses for $\epsilon = 1\%$

Target return (%)	WVaR	NVaR	ARVaR	CVaR-based VaR	ESVaR
0.300	6.351	6.351	4.279	4.482	4.342
0.500	9.729	9.720	3.643	7.237	5.591
0.600	11.954	11.948	3.325	8.491	4.513
0.800	16.404	16.390	3.761	8.857	5.249
1.000	20.854	20.838	4.752	8.529	7.309
1.200	25.304	25.294	5.790	9.962	7.066
1.400	29.754	29.743	6.828	11.622	7.870
1.600	34.204	34.183	7.866	13.523	8.808
2.000	43.104	43.089	9.942	16.403	11.780
2.500	54.224	54.240	12.536	19.573	13.864
3.000	65.353	65.373	15.201	22.633	19.954

Figure 6 Realized Cumulative Portfolio Return for Out-of-Sample Observations When a Simple Buy-and-Hold Strategy is Employed



Note. The target return used to obtain the portfolio weights is 2.5%, and $\epsilon = 1\%$.

7. Concluding Remarks

We present a distribution-parameter-based approach to VaR optimization based on robust optimization techniques and the concept of asymmetric variability measures introduced in Chen et al. (2007). We show that the resulting portfolio risk measure is coherent, and that it approximates another coherent risk measure that has been gaining in popularity in the literature and in industry—the Conditional VaR. We also studied how the Asymmetry-Robust VaR approach performs in controlled and real-data experiments relative to other distribution-parameter-based and sample-based VaR optimization approaches.

Sample-based approaches for VaR optimization take into consideration every observation in the sample. In some cases, however, it may be preferable to use a more general estimate of the variability of a distribution, either because of insufficient data, or to avoid overfitting. We believe that in such instances the idea of the analytical, less data-set-specific Asymmetry-Robust VaR approach holds promise. Our computational experiments support this belief, and suggest that in cases when the underlying distributions are skewed, the ARVaR approach may be preferable also to distribution-parameter-based VaR optimization approaches such as the Normal VaR that use second-moment information.

Acknowledgments

The research of the first author was partially supported by Singapore-MIT Alliance, NUS Risk Management Institute, and NUS Start-up Grants R-146-050-070-133 and R146-050-070-101. The research of the second author was supported by the Gill grant from the Babson College Board of

Research. The research of the third author was partially supported by Singapore-MIT Alliance, NUS Risk Management Institute, and NUS Academic Research Grants R-314-000-066-122 and R-314-000-068-122.

References

- Artzner, P., F. Delbaen, J. M. Eber, D. Heath. 1997. Thinking coherently. *Risk* 10 68–71.
- Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Math. Oper. Res.* 23(4) 769–805.
- Ben-Tal, A., A. Nemirovski. 1999. Robust solutions to uncertain programs. *Oper. Res. Lett.* 25 1–13.
- Ben-Tal, A., A. Nemirovski. 2001. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadelphia.
- Bertsimas, D., M. Sim. 2004. The price of robustness. *Oper. Res.* 52(1) 35–53.
- Bertsimas, D., D. Pachamanova, M. Sim. 2004. Robust linear optimization under general norms. *Oper. Res. Lett.* 32 510–516.
- Cambell, J., A. Lo, A. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.
- Cariño, D. R., A. L. Turner. 1998. Multiperiod asset allocation with derivative assets. W. T. Ziemba, J. M. Mulvey, eds. *Worldwide Asset and Liability Modeling*. Cambridge University Press, Cambridge, MA, 182–204.
- Chen, W., M. Sim. 2006. Goal driven optimization. *Oper. Res.* Forthcoming.
- Chen, X., M. Sim, P. Sun. 2007. A robust optimization perspective on stochastic programming. *Oper. Res.* 55(6) 1058–1071.
- Dowd, K. 1998. *Beyond Value at Risk: The New Science of Risk Management*. John Wiley & Sons, New York.
- Duffee, G. 2002. The long-run behavior of firms' stock returns: Evidence and interpretations. Working paper, Haas School of Business, University of California at Berkeley, Berkeley, CA.
- El Ghaoui, L., M. Oks, F. Oustry. 2003. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Oper. Res.* 51(4) 543–556.
- El Ghaoui, L., F. Oustry, H. Lebret. 1998. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.* 9(1) 33–52.
- Fama, E. 1976. *Foundations of Finance*. Basic Books, New York.
- Föllmer, H., A. Schied. 2002. Robust preferences and convex measures of risk. K. Sandmann, P. Schönbucher, eds. *Advances in Finance and Stochastics*. Springer-Verlag, Berlin, 39–56.
- Jorion, P. 2000. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, New York.
- Konno, H., H. Yamazaki. 1991. Mean-absolute deviation portfolio optimization model and its applications to the Tokyo stock market. *Management Sci.* 37(5) 519–531.
- Larsen, N., H. Mausser, S. Uryasev. 2002. Algorithms for optimization of value-at-risk. P. Pardalos, V. K. Tsitsiringos, eds. *Financial Engineering, e-Commerce and Supply Chain*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 129–157.
- Levy, H. 1992. Stochastic dominance and expected utility: Survey and analysis. *Management Sci.* 38(4) 555–593.
- Markowitz, H. M. 1952. Portfolio selection. *J. Finance* 7 77–91.
- Rockafellar, R. T., S. Uryasev. 2000. Optimization of conditional value-at-risk. *J. Risk* 2(3) 21–41.
- Rockafellar, R. T., S. Uryasev. 2002. Conditional value-at-risk for general loss distributions. *J. Banking Finance* 26(7) 1443–1471.
- Schönbucher, P. 2000. Factor models for portfolio credit risk. Working paper, University of Bonn, Germany. <http://www.gloriamundi.org>.
- Yamai, Y., T. Yoshida. 2002. On the validity of value-at-risk: Comparative analyses with expected shortfall. *Monetary Econom. Stud.* 20(1) 57–85.